RESEARCH

Confirmation of Molecular Markers and Agronomic Traits Associated with Seed Phytate Content in Two Soybean RIL Populations

A.M. Scaboo,* V.R. Pantalone, D.R. Walker, H.R. Boerma, D.R. West, F.R. Walker, and C.E. Sams

ABSTRACT

The concentration of phytate is an important consideration when analyzing feed grain for livestock. Simple sequence repeat markers Satt237 and Satt561 were found to be linked to quantitative trait loci (QTLs) for phytate concentration in soybean [*Glycine max* (L.) Merr.] seed. Previous research also established a significant correlation between seed inorganic phosphorus (Pi) and seed phytate. The objectives of this study were to confirm these QTL in independent environments and populations, to compare the effectiveness of marker-assisted selection with that of phenotypic selection for low phytate, and to evaluate agronomic traits associated with low phytate soybean. The low phytate soybean mutant CX1834-1-2 was crossed with the cultivar 5601T and line S97-1688 to form two populations. The Satt237 marker was linked to a major QTL ($R^2 > 0.10$) governing seed phytate, and Satt561 was linked to a minor QTL ($R^2 < 0.10$) in both populations. These two confirmed QTL have been named cqPha-001 and cqPha-002. The phenotypic correlations were significant ($P < 0.05$) for plant height ($r = 0.26$), seed protein ($r = -0.19$), and seed oil ($r = 0.17$). Phenotypic selection was found to be more effective in lowering mean phytate levels and more cost efficient.

A.M. Scaboo, Dep. of Crop, Soil, and Environmental Sciences, Univ. of Arkansas, 115 Plant Science Bldg., Fayetteville, AR 72701; V.R. Pantalone, D.R. West, and C.E. Sams, Dep. of Plant Sciences, Univ. of Tennessee, 2431 Joe Johnson Dr., 252 Ellington Plant Science Bldg., Knoxville, TN 37996; F.R. Walker, Dep. of Biosystems Engineering and Soil Science, Univ. of Tennessee, 2506 E J Chapman Dr., Knoxville, TN 37996; D.R. Walker, USDA-ARS Soybean/Maize Germplasm, Pathology, and Genetics Research Unit, Urbana, IL 61801; H.R. Boerma, Dep. of Crop and Soil Sciences, Univ. of Georgia, Center for Applied Genetic Technologies, 111 Riverbend Rd., Athens, GA 30602. Received 8 Nov. 2007. *Corresponding author (ascaboo@uark.edu).

Abbreviations: ETREC, East Tennessee Research and Education Center; LG, linkage group; MAS, marker-assisted selection; MG, maturity group; PCR, polymerase chain reaction; Pi, inorganic phosphorus; QTL, quantitative trait locus; RIL, recombinant inbred lines; SSR, simple sequence repeat.

The concentration of phytate, a mixed cation salt of phytic acid (myo-inositol 1,2,3,4,5,6-hexakisphosphate), is an important consideration when analyzing grain for livestock feed rations. The phosphorus found in phytate is mostly unavailable to humans and monogastric animals such as poultry, swine, and fish, and it binds valuable mineral nutrients including calcium, magnesium, iron, potassium, and zinc into an unavailable form (Erdman, 1979; Raboy, 2002). Feed rations for nonruminant animals currently include phytase, an enzyme responsible for the dephosphorylation of phytate, and/or inorganic phosphorus (Pi) supplements (Lei and Porres, 2003). Improper application of manure from swine and poultry production to the soil can result in nonpoint source Pi contamination of ground or surface waters. Such contamination can cause eutrophication in aquatic ecosystems, promoting algal

Published in Crop Sci. 49:426–432 (2009).
doi: 10.2135/cropsci2007.11.0614

677 S. Segoe Rd., Madison, WI 53711 USA

blooms and reducing available oxygen for macro-organisms (Daniel et al., 1994; Ertl et al., 1998). The development of soybean [*Glycine max* (L.) Merr.] cultivars that are lower in seed phytate should improve availability of mineral nutrients to growing monogastric animals, while at the same time reducing damage to ecosystems inundated with non-point source Pi pollution (Powers et al., 2007).

Wilcox et al. (2000) successfully developed mutant soybean lines with ~1.9 g kg^{-1} of phytate P and ~3.1 g kg^{-1} of Pi, compared to normal conventional soybean cultivars that had ~4.3 g kg^{-1} of phytate P and ~0.7 g kg^{-1} of Pi. One of these lines, CX1834-1-2, was used as the low phytate male donor parent in populations developed by the University of Tennessee and the University of Georgia to identify molecular markers associated with this trait (Walker et al., 2006). Using simple sequence repeat (SSR) markers, Walker et al. (2006) found two genomic regions associated with SSR markers that correlated with seed phytate content in soybean. On linkage group (LG) N, Satt237 was found to be linked to a major locus associated with seed phytate, and on LG L, Satt527 and Satt561 were found to be linked to a locus with a smaller effect on seed phytate concentration. On the basis of interval mapping LOD score plots, Satt237 and Satt561 each appeared to be within a few centimorgans of the phytic acid quantitative trait loci (QTLs) on their respective linkage groups. The molecular data together with phenotypic data for the eight different genotypic classes supported the conclusion of Oltmans et al. (2004) that the low phytate phenotype of the CX1834-1–derived lines results from duplicate dominant epistasis at two independent loci. Simple sequence repeat markers Satt237 and Satt561 were used in this study to confirm the association of these genomic regions with the low phytate phenotype in genetic backgrounds and environments different from those of the original mapping populations.

The confirmation and validation of reported QTLs requires the identification of the same QTL in parallel and independent populations and environments, thus QTL confirmation increases the likelihood that marker-assisted selection (MAS) will be successful and demonstrates that the marker associated with a QTL is effective for selection irrespective of genetic background and environment (Fasoula et al., 2004). The objectives of this study were to confirm two QTLs associated with seed phytate concentration in different environments and parallel segregating soybean populations, to compare the effectiveness of MAS vs. phenotypic selection for low phytate, and to evaluate agronomic traits associated with low phytate soybean.

MATERIALS AND METHODS

Population Development and Experimental Field Procedures

A low-phytate, maturity group (MG) III germplasm, CX1834-1-2, that was developed by Dr. J.R. Wilcox of the USDA-ARS at Purdue University was crossed with a high yielding MG V cultivar, 5601T (Pantalone et al., 2003). A second population was developed by crossing CX1834-1-2 with a competitive-yielding MG V line, S97-1688 (Anand et al., 2004), which has broad soybean cyst nematode (*Heterodera glycines* Ichinohe) resistance. The crosses were made in the summer of 2000 at the Plant Science Unit of the East Tennessee Research and Education Center (ETREC) in Knoxville, TN, and the following description of the generational advancement applies to both populations. Approximately 20 F_1 seeds were harvested in the fall of 2000 and sent to Costa Rica for generation advancement to the F_5 stage via single seed descent (Fehr, 1991). The F_1 plants were grown as a single row and bulk harvested during the spring of 2001. Approximately 800 F_2 seeds were row planted in Costa Rica during the spring of 2001, and single seeds were harvested from each F_2 plant and bulked. The procedure of harvesting a single seed from each plant and bulking was followed for the F_3 and F_4 generations, with plants in each generation growing for ~4 mo in Costa Rica. Approximately 700 F_5 seeds were then shipped back to Tennessee for planting during the 2002 growing season. F_5 single plants were grown in Knoxville, TN, at the Plant Science Unit of the ETREC in eight 6.0-m rows during the summer of 2002. Approximately 700 F_5 single plants were harvested and threshed during the fall of 2002, and were sent to Costa Rica during the winter for seed increase. In the summer of 2003, 313 $F_{5:7}$ recombinant inbred lines (RILs) from the 5601T × CX1834-1-2 population were grown in 3.0-m one-row plots, replicated three times, at the Plant Science Unit of the ETREC in Knoxville, TN. A total of 815 $F_{5:7}$ RILs from the cross S97-1688 × CX1834-1-2 were also grown in 3.0-m one-row plots, unreplicated, at the Holston Unit of the ETREC during the summer of 2003. A total of 187 RILs were randomly chosen from the 5601T × CX1834-1-2 population and divided into three groups. This was strictly for ease of harvesting in 2004, and was based on maturity observations made in the field during 2003. In 2004, the 187 $F_{5:8}$ RILs from the 5601T × CX1834-1-2 population were planted in two-row 6.1-m long plots in a randomized complete block design with three replications at three locations: the Plant Science Unit of the ETREC in Knoxville, TN (Sequatchie silt loam soil [fine-loamy, siliceous, semiactive, thermic Humic Hapludult]); the Research and Education Center at Milan, TN (Calloway silt loam soil [fine-silty, mixed, active, thermic Aquic Fraglossudalf]); and the Highland Rim Research and Education Center in Springfield, TN (Baxter Cherty silt loam [fine, mixed, semiactive, mesic Typic Paleudalf]). Rows were end-trimmed to a length of 4.9 m before harvest. A total of 186 $F_{5:7}$ RILs were randomly chosen from the S97-1688 × CX1834-1-2 population and were grown as 6.1-m long single-row plots in an unreplicated field study at the Plant Science Unit of the ETREC in Knoxville, TN, during 2004. The following phenotypic traits were measured in 2004 by observations in the field for the 5601T × CX1834-1-2 population: flower color (purple, white, or segregating), maturity (days from planting to physiological maturity), lodging (scored from 1 = erect to 5 = prostrate), seed yield (kg ha^{-1}), and plant height (cm).

DNA Isolation and Polymerase Chain Reaction (PCR)

In 2003, approximately 7 to 10 leaves of each $F_{5:7}$ RIL were sampled from one field replication, in both populations, for DNA extraction. Samples were taken randomly from each row

and stored at −80°C until the extraction procedure was performed. DNA from these leaf samples was isolated using a Qiagen Plant DNeasy Extraction Kit (Qiagen, Valencia, CA).

Progeny from both RIL populations were screened with the SSR markers Satt237 (LG N) and Satt561 (LG L) developed by Cregan et al. (1999). Polymerase chain reaction amplifications were performed in a 96-well MBS Hybaid thermocycler (Hybaid, Franklin, MA). The 12.5-μL PCR mix consisted of 2.5 μL of 10 ng μL^{-1} template DNA, 3.8 μL double distilled H_2O, 1.3 μL of 25 μM forward D4 labeled primer (synthesized by Proligo, Boulder, CO) and 25 μM reverse primer (synthesized by Sigma Genosys, Woodlands, TX), and 5.0 μL of Hotmaster *Taq* (Eppendorf Brinkman Instruments, Westbury, NY). The amplification conditions consisted of 94°C for 5 min followed by 35 cycles of 94°C denaturation for 25 s, 57°C annealing for 30 s, and 72°C extension for 30 s, ending with a final cycle of 72°C for 5 min. The PCR products were separated using capillary electrophoresis in a Beckman CEQ 8000 (Fullerton, CA). Fragment lengths were determined using Beckman CEQ Fragment Analysis Software. Satt237 (LG N) amplicons were 257 and 239 bp long for CX1834-1-2 and 5601T alleles, respectively, and Satt561 (LG L) amplicons were 266 and 260 bp in length for CX1834-1-2 and 5601T alleles, respectively.

The extracted DNA from 55 progeny lines in the S97-1688 × CX1834-1-2 population was taken to the University of Georgia (Athens, GA) for SSR marker Satt561 screening. Polymerase chain reaction and polyacrylamide electrophoresis protocols were the same as those described by Li et al. (2001), except that a PTC-225 DNA Engine Tetrad (MJ Research, Waltham, MA) thermalcycler was used. Polymerase chain reaction amplicons were separated by polyacrylamide gel electrophoresis on an ABI 377 automated DNA sequencer (Applied Biosystems, Foster City, CA). The RIL progeny were scored at the two marker loci as follows: 1 = homozygous for the 5601T or S97-1688 allele, 2 = heterozygous, and 3 = homozygous for the CX1834-1-2 allele.

Inorganic Phosphorus and Phytate Composition Analysis

Inorganic phosphorous concentrations were determined using a modified version of a colorimetric assay developed by Raboy et al. (2000), which was an adaptation of the assay described by Chen et al. (1956). Approximately 30 g of soybean seeds, randomly chosen from F_9 seed harvested in 2004 from the $F_{5:8}$ RIL plots, were ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, S-26321, Hogana, Sweden) for 20 s. This setting produced ground soybean seed powder with a uniform particle size. Approximately 100 mg of ground soybean seed tissue was placed into a 1.5-mL microcentrifuge tube for Pi extraction, and the weight per sample was recorded for calculation of Pi on a dry weight basis. A 1.0-mL aliquot (10 μL per mg of seed tissue) of the extraction buffer (12.5% trichloroacetic acid and 25 mM $MgCl_2$) was added to each sample, vortexed, and allowed to incubate overnight (~16 h). The samples were then vortexed and allowed to settle for 5 min before aliquots of extraction solution were transferred to 96-well storage plates, and then centrifuged at 604 *g* for 3 min.

Three 10-μL subsamples were taken from each extraction sample and then tested in the same 96-well assay plate to obtain a mean estimate of Pi levels for each RIL. The 10-μL subsamples were diluted into 90 μL of double distilled H_2O (100 μL total volume) and mixed with 100 μL of Chen's Reagent, which consists of 1 volume of 3 M H_2SO_4, 1 volume of 0.02 M ammonium molybdate, 1 volume of 10% ascorbic acid, and 2 volumes of water (Chen et al., 1956). The reaction was allowed to proceed at ambient temperature for exactly 1 h, after which Pi concentrations were determined using a BioTek PowerWave XS Microplate Spectrophotometer (BioTek Instruments, Winooski, VT) set at a wavelength of 882 nm. Eight standards, representing 0, 0.155, 0.465, 0.930, and 1.395 μg of P, based on the atomic weight of P, were used to obtain a standard curve, from which sample concentrations were estimated. In addition to the five standards used by Raboy et al. (2000), standards representing 1.860, 2.320, and 2.640 μg P were also included. This made it possible to use the original sample dilution ratio (10 μL extract solution + 90 μL H_2O) described by Raboy et al. (2000). Direct measurements of seed phytate concentrations were conducted at the USDA-ARS in Raleigh, NC, with binary high pressure liquid chromatography using a 50 by 4.6 mm PL-SAX strong anion exchange column (Polymer Laboratories, Amherst, MA) equipped with a 7.5 by 4.6 mm guard column. Ground soybean seed tissue was analyzed for phytic acid concentration from a representative sample (85 RILs of the 5601T × CX1834-1-2 population) as described by Israel et al. (2006).

Protein and Oil Analyses

Approximately 30 g of soybean seeds, randomly chosen from F_9 seed harvested in 2004 from the $F_{5:8}$ RIL plots were ground in a water-cooled Knifetec 1095 Sample Mill (FOSS Tecator, Hogana, Sweden) for 20 s. Protein and oil concentrations (dry weight) were obtained using a NIRS instrument (NIRS 6500; FOSS North America, Eden Prairie, MN) according to the user's manual. Diagnostics tests ensured that the instrument passed three different tests for instrument response, wavelength accuracy, and NIRS repeatability (Panthee et al., 2004). A room dehumidifier was used throughout the analysis to maintain the humidity at 40%, and the room temperature was approximately 20°C. Ground soybean samples were scanned using Winisi II 1.5 software (FOSS North America). The instrument was left on for the entire period of analysis, and diagnostics were performed every day until scanning of all samples was completed.

Data Analyses

Associations between molecular data and phenotypic data used to validate QTLs were analyzed using single factor analysis of variance with the PROC GLM procedure of SAS software v9.0 (SAS Institute, 2003). The few RILs that were heterozygous at the Satt237 or Satt561 loci were excluded from the analysis because the RILs were F_5–derived, and because our primary interest was to confirm the additive genetic effect of QTLs from lines homozygous at those two loci. Mean seed Pi concentrations averaged across all replicates and locations were analyzed with the PROC MEANS procedure in SAS software v9.0 (SAS Institute, 2003). The PROC GLM procedure also was used to determine the analysis of variance for Pi, with location and replication as random blocking factors. Heritability of the phenotypic traits in this population were estimated on an entry mean and plot basis (Fehr, 1991; Nyquist, 1991). The GLM procedure of SAS was used to estimate variance components for calculating heritability estimates. Mean phenotypic correlations, averaged

across replications and locations were determined using the PROC CORR procedure in SAS v9.0 (SAS Institute, 2003). Mean phenotypic traits averaged across all replicates and locations were analyzed with the PROC MEANS procedure in SAS software v9.0 (SAS Institute, 2003). The PROC GLM procedure also was used to determine the analysis of variance for all other measured traits, with location and replication considered random effects.

RESULTS AND DISCUSSION

Although the inverse relationship between seed phytate and seed Pi in the mutant plant from which CX1834-1-2 was derived has been previously reported (Wilcox et al., 2000; Oltmans et al., 2005; Spear and Fehr, 2007), it was important to confirm this in the genetic backgrounds used for this research before relying entirely on assays for seed Pi concentration to confirm low phytate QTLs. This relationship is important because measuring Pi levels using an assay described by Raboy et al. (2000) is substantially easier, quicker, and less expensive than measuring phytate. This is a modification of the colorimetric assay developed by Chen et al. (1956). There was a significant linear regression between seed Pi and seed phytate concentrations (Fig. 1), and the variation in seed Pi explained ~84% of the variation in seed phytate. As a result, Pi was considered a reliable indicator of the relative level of phytate in soybean seed.

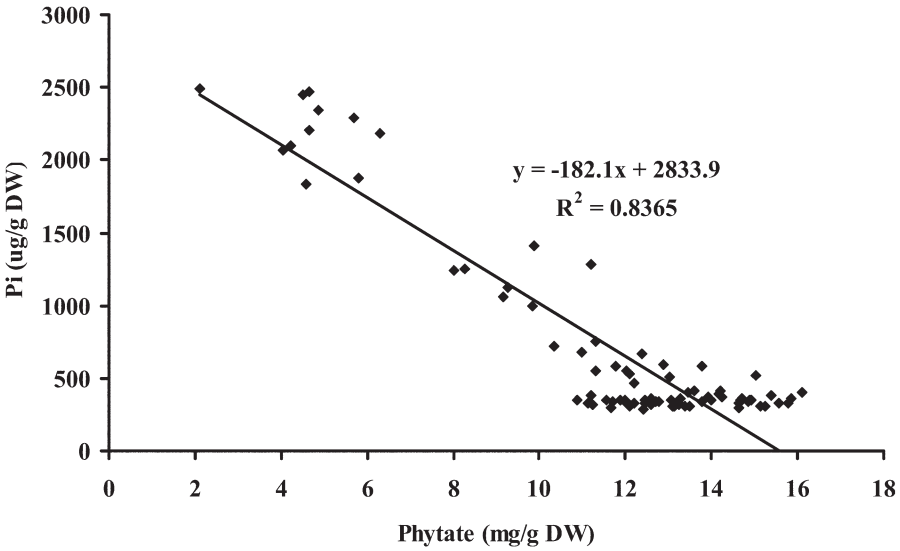

Figure 1. Scatter plot and linear regression for seed phytate and inorganic phosphorus (Pi) concentrations of 85 $F_{5:7}$ recombinant inbred lines from the cross 5601T × CX1834-1-2.

Table 1. Quantitative trait loci for seed inorganic phosphorus (Pi) concentration in 187 $F_{5:8}$ recombinant inbred lines of 5601T × CX1834-1-2 grown in 2004 at three environments with three replications, using single factor analysis of variance.

SSR locus	Combined				Environment†					
	Allelic mean Pi‡		*P*	R^2	HRREC		ETREC		MREC	
	56/56	CX/CX			*P*	R^2	*P*	R^2	*P*	R^2
	μg Pi g^{-1}			%		%		%		%
Satt237	441	1196	<0.0001	30.2	<0.0001	30.7	<0.0001	28.9	<0.0001	32.7
Satt561	557	976	<0.0001	8.2	<0.0001	8.6	<0.0001	9.1	<0.0001	7.6
DLR§	394	1678	<0.0001	56.4	<0.0001	56.1	<0.0001	56.5	<0.0001	57.1

†HRREC, Highland Rim Research and Education Center at Springfield, TN; ETREC, East Tennessee Research and Education Center at Knoxville, TN; MREC, Milan Research and Education Center at Milan, TN.

‡56/56, homozygous 5601T marker allele; CX/CX, homozygous CX1834-1-2 allele.

§Dual Locus Regression, including the Satt237 × Satt561 interaction.

For the 5601T × CX1834-1-2 population, Satt237 and Satt561 were confirmed to be associated with a QTL ($P < 0.0001$) for seed Pi in all environments, and in the combined data set (Table 1). Satt237 was confirmed to be linked to a major QTL ($R^2 = 30.2\%$) on LG N, and Satt561 was found to be linked to a minor QTL ($R^2 = 8.2\%$) on LG L. These data show that there is a two- to fourfold difference in the mean Pi value between lines with the SSR allele from the normal phytate parent and lines with the SSR allele from CX1834-1-2. For the S97-1688 × CX1834-1-2 population, Satt237 was linked to a major QTL ($P < 0.0001$, $R^2 = 10.1\%$) and Satt561 was linked to a minor QTL ($P = 0.0172$, $R^2 = 3.5\%$) (Table 2). The mean seed Pi concentration associated with normal and low phytate RILs for each SSR marker genotypic class in this population showed a 0.5- to 4-fold difference. The Satt237 × Satt561 interaction was also significant ($P < 0.0001$) in the multiple regression analysis for both populations. These data confirm that the QTL identified by Walker et al. (2006) on LG N that is linked to Satt237 and the one on LG L that is linked with Satt561 are associated with seed Pi concentration. These two confirmed QTLs have been named cqPha-001 and cqPha-002 in accordance with the nomenclature guidelines of the Soybean Genetics Committee.

The genotypic class with the highest Pi levels in both populations was the group that was homozygous for the CX1834-1-2 allele at both the Satt237 and Satt561 loci

Table 2. Quantitative trait loci for seed inorganic phosphorus (Pi) concentration in 186 $F_{5:8}$ recombinant inbred lines of S97-1688 × CX1834-1-2 grown in 2004 at Knoxville, TN, using single factor analysis of variance.

SSR locus†	Allelic mean Pi‡		*P*	R^2
	S97/S97	CX/CX		
	—— μg Pi g^{-1} ——			
Satt237	431	795	<0.0001	10.1
Satt561	408	652	0.0172	3.5
DLR§	373	1464	<0.0001	28.3

†SSR, simple sequence repeat.

‡S97/S97, homozygous S97-1688 allele; CX/CX, homozygous CX1834-1-2 allele.

§Dual Locus Regression, including the Satt237 × Satt561 interaction.

(Fig. 2 and 3). These data show that the lines with the highest Pi concentrations (and lowest phytate concentrations) were those that carried the SSR marker alleles inherited from the low phytate parent, CX1834-1-2. In contrast, the lines with the lowest concentration of Pi were those that were homozygous for the SSR marker alleles inherited by the normal phytate parents. The efficiency of SSR markers to select superior (low-phytate) progenies can be described when observing mean Pi concentrations in both populations. The mean Pi concentration for the 5601T × CX1834-1-2 population was 788 μg g^{-1} (combined data), and the mean Pi concentration for the S97-1688 × CX1834-1-2 population was 561 μg g^{-1}. If only individuals with the CX1834-1-2 allele at Satt237 are selected from the 5601T × CX1834-1-2 population, the mean Pi concentration increases to 1196 μg g^{-1} (Table 1). When both SSR markers are used for the selection of progeny homozygous for alleles from CX1834-1-2 in that population, the mean Pi concentration is more than doubled, from 788 μg g^{-1} to 1679 μg g^{-1}. When both SSR markers are used for the selection of homozygous progeny in the S97-1688 × CX1834-1-2 population, the mean Pi concentration is also more than doubled, from 561 to 1464 μg g^{-1} (Table 2). The efficiency of using these SSR markers to select superior lines can also be observed by determining the percentage of lines that were identified as low phytate lines after MAS. Table 3 shows the percentage of individuals that were correctly selected (i.e., individuals that inherited the CX1834-1-2 allele at the Satt237 and/or Satt561 loci and also expressed a low phytate phenotype) on the basis of seed Pi concentration. The percentage of individuals with a low phytate phenotype ranged between 9.7 and 25.0% after MAS using only one of the two SSR markers, depending on which marker is chosen and which breeding population that marker is applied to. After using both SSR markers for selection, only 50% of the RILs expressed the low phytate phenotype in both populations. These results show that MAS selection for low phytate, when using SSR markers Satt237 and Satt561, is only 50% effective in selecting low phytate individuals when compared to phenotypic selection in RIL populations. This relatively low percentage may be a result of the loss of low phytate individuals during population development due to the correlation between the low phytate phenotype and reduced seedling emergence that was reported by Meis et al. (2003). Although MAS was not as effective in identifying low phytate individuals as phenotypic selection, all low phytate individuals inherited the CX1834-1-2 form of each of the Satt237 and Satt561 SSR marker loci.

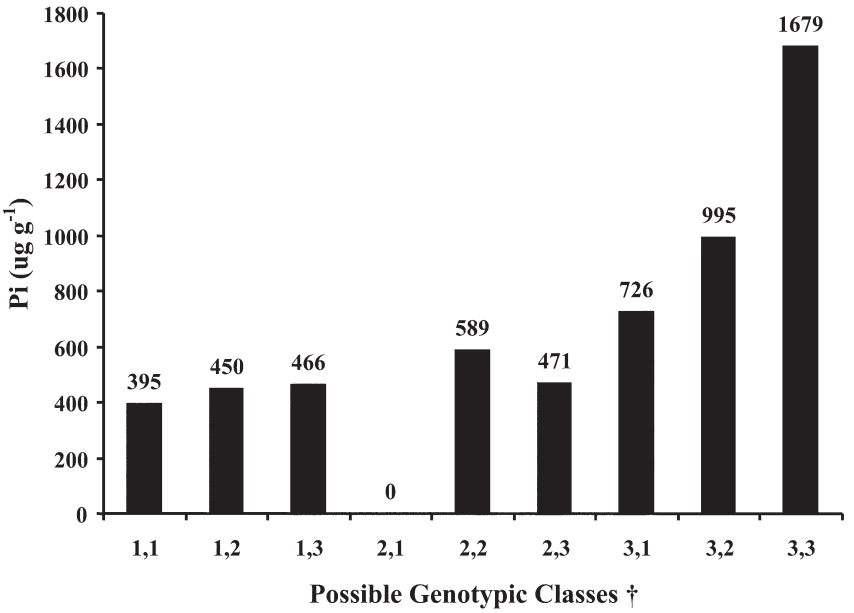

Figure 2. Mean soybean seed inorganic phosphorus (Pi) concentration for all possible genotypic classes, based on simple sequence repeat markers Satt237 and Satt561, for 187 $F_{5:8}$ recombinant inbred lines grown in three environments in 2004 from the 5601T × CX1834-1-2 population. †1 = Homozygous for the S97-1688 locus, 2 = Heterozygous locus, 3 = Homozygous for the CX1834-1-2 locus (Satt237 score is first number, Satt561 score is second number).

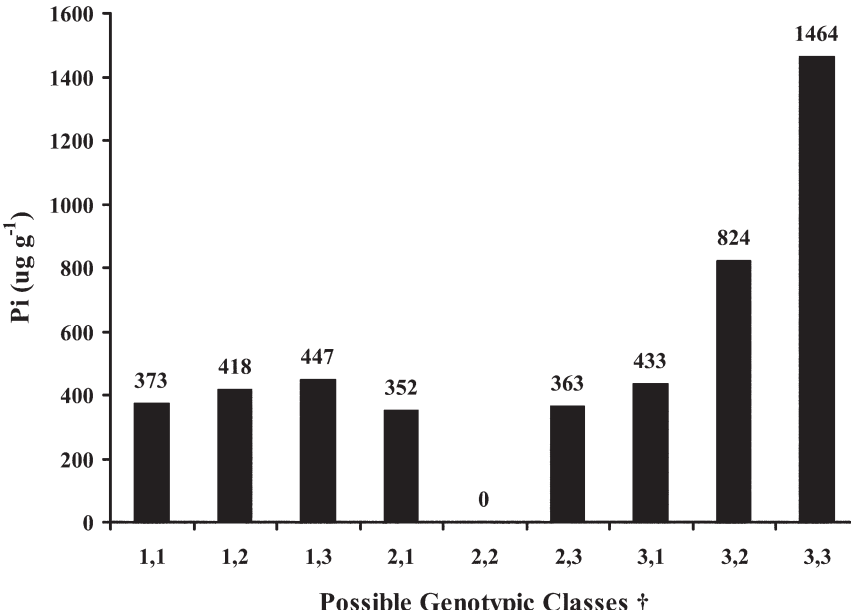

Figure 3. Mean soybean seed inorganic phosphorus (Pi) concentration for all possible genotypic classes, based on simple sequence repeat markers Satt237 and Satt561, for 186 $F_{5:8}$ recombinant inbred lines grown in 2004 from the S97-1688 × CX1834-1-2 population. †1 = Homozygous for the S97-1688 locus, 2 = Heterozygous locus, 3 = Homozygous for the CX1834-1-2 locus (Satt237 score is first number, Satt561 score is second number).

This observation leads to the conclusion that a soybean line that possesses the CX1834-1-2 form of the SSR marker allele would have the highest probability of producing a low phytate phenotype, and that recombination between the QTLs and their linked SSR markers is not the reason for the inefficiency of MAS. Thus, MAS with SSR markers Satt237 and Satt561 will capture all available low phytate individuals, but simply inheriting the CX1834-1-2 allelic form of SSR markers Satt237 and Satt561 is not necessarily indicative of a low phytate phenotype. These markers may be most useful compared to phenotypic identification for distinguishing heterozygous progeny in backcross populations. The introgression of the low phytate phenotype into prominent cultivars using a backcrossing technique would require the identification of plants that are heterozygous for the two QTLs over multiple cycles of backcrossing. Figures 2 and 3 show that heterozygous individuals were similar in Pi concentration to individuals homozygous at both alleles for the normal phytate parent. This would make it extremely difficult to phenotypically distinguish individuals carrying the desirable alleles from those homozygous for one or both of the recurrent parent alleles, particularly since most breeders would select visually using the Raboy et al. (2000) seed Pi assay, rather than using a spectrophotometer. In addition, breeders would be able to select plants for the next round of backcrossing before flowering and would also have the option of selecting for other important QTLs.

Table 3. Efficiency of marker-assisted selection, as the percentage of progeny recombinant inbred lines that were correctly identified as low phytate lines using simple sequence repeat markers Satt237 and Satt561, in two soybean populations and combined.

Population	Allelic frequency†			LPL‡	% Correctly selected		
	Satt237	Satt561	Both§		Satt237	Satt561	Both§
5601T × CX1834-1-2	80	94	40	20	25.0	21.3	50.0
S97-1688 × CX1834-1-2	56	103	19	10	17.9	9.7	52.6
Combined	136	197	59	30	22.1	15.2	50.8

†Frequency of individuals with homozygous CX1834-1-2 allele.

‡LPL, low phytate lines.

§Frequency of lines with homozygous CX1834-1-2 allele for both Satt237 and Satt561.

Table 4. Descriptive statistics for agronomic and seed quality traits in 187 $F_{5:8}$ recombinant inbred lines from the cross 5601T × CX1834-1-2, averaged over three replications and three environments in 2004.

Trait	Min.	Mean	Max.	SD	5601T	CX1834-1-3	h^{2}‡
Yield, kg ha^{-1}	1363	2372	3178	5.3	3251	1236	0.49, 0.16
Height, cm	58.5	83.5	113	4.9	83.8	59.8	0.89, 0.61
Lodging†	1.3	1.9	3.6	0.5	1.5	1.8	0.77, 0.36
Maturity, day	251	266	289	7.0	278	251	0.72, 0.24
Pi, μg g^{-1}	348	788	3123	650.0	394	2137	0.97, 0.92
Protein, g kg^{-1}	391	426	465	17.3	412	437	0.95, 0.78
Oil, g kg^{-1}	149	177	198	11.1	180	182	0.95, 0.79

†Lodging scored from 1 to 5 scale, 1 = all plants erect and 5 = all plants prostrate.

‡Heritability estimated on an entry mean and plot basis.

Descriptive statistics for the agronomic and seed quality traits for RILs measured in this study are presented in Table 4. CX1834-1-3, a sister line of the parental line CX1834-1-2, was used as a check line in this field study in 2004 (due to limited seed availability). Heritability was estimated for each trait on an entry means and plot basis (Table 4). The plot and entry mean based heritability estimates for seed Pi were high, suggesting that most of the phenotypic variation in seed Pi is due to genes inherited from the parental lines. This high heritability estimate for seed Pi also suggests that the inheritance of the low phytate phenotype is largely genetically controlled, so molecular identification of all genes involved should be readily accomplished. Phenotypic correlation coefficients for agronomic and seed quality traits are shown in Table 5. The phenotypic correlation between seed Pi concentration and yield was not significant, which is consistent with the studies of Oltmans et al. (2005), Hulke et al. (2004), and Spear and Fehr (2007) that showed the low phytate trait in soybean does not negatively impact seed yield. Inorganic phosphorus also was not significantly correlated with maturity and lodging. Significant phenotypic correlations were found between Pi and plant height (0.26), seed protein concentration (−0.19), and seed oil concentration (0.17). Although these correlations were significant, they are sufficiently low to permit the selection of acceptable low phytate lines for each trait. These data confirm data presented by Oltmans et al. (2005) that showed that there was no significant difference in normal and low phytate lines for the measured agronomic traits, and that breeders

Table 5. Phenotypic correlation coefficients among agronomic and seed quality traits in 187 $F_{5:8}$ recombinant inbred lines from the cross 5601T × CX1834-1-2, averaged across three replications and three environments in 2004.

Trait	Height	Lodging	Yield	Inorganic P	Protein	Oil
Maturity	0.45†	NS‡	0.36†	NS	−0.35†	0.44†
Height		0.35†	0.22**	0.26†	NS	0.24**
Lodging§			NS	NS	NS	NS
Yield				NS	NS	0.21**
Inorganic P					−0.19*	0.17*
Protein						−0.80†

*Significant at the 0.05 probability level.

**Significant at the 0.01 probability level.

†Significant at the 0.0001 probability level.

‡NS, not significant.

§Lodging scored on a 1 to 5 scale; 1 = all plants erect and 5 = all plants prostrate.

should be able to develop low phytate soybean lines with normal or desirable agronomic traits.

The efficiency of MAS for the low phytate soybean phenotype must also be understood from an economical standpoint because resources are of major concern when determining a selection strategy for a new trait in a breeding program. At the University of Tennessee we estimated the price of phenotypic selection per 96 individuals to be US $6.33, requiring a total of ~23 person hours. This is compared to the relatively high cost of US $250.00 for genotyping 96 individuals, requiring ~48 person hours. These results suggest that phenotyping for low phytate soybeans is likely to be more effective and cost efficient than MAS with SSR markers Satt237 and Satt561 for identifying plants that are homozygous for the low phytate allele at the phytate loci on LGs L and N. However, the use of the molecular markers in a backcrossing program would be useful because individuals heterozygous for the two alleles cannot be determined by phenotypic assays, and because a breeder would like to identify the heterozygotes from the previous backcross generation before making the next round of pollinations. This is primarily due to the duplicate dominant epistasis described by Oltmans et al. (2004) and confirmed with molecular data from Walker et al. (2006) and from the present study. Furthermore, the increasing availability of single nucleotide polymorphism markers that could be used for MAS and the potential to reduce cost of DNA extraction, currently a large part of the total cost of MAS, are likely to substantially reduce the cost of MAS in the future as well as provide finer mapping techniques for identifying prominent QTLs.

Acknowledgments

Funding from the United Soybean Board, Better Bean Initiative, The Consortium for Plant Biotechnology Research, and by the Tennessee Soybean Promotion Board is greatly appreciated. Continued support by the Tennessee Agriculture Experiment Station is also highly appreciated. We would like to express our thanks to Debbie Ellis for her contributions to this project.

References

Anand, S.C., J.G. Shannon, J.A. Wrather, P.R. Arelli, D.A. Sleper, and L.D. Young. 2004. Registration of S97-1688 soybean germplasm line. Crop Sci. 44:698–700.

Chen, P.S., Jr., T.Y. Toribara, and H. Warner. 1956. Microdetermination of phosphorus. Anal. Chem. 28:1756–1758.

Cregan, P.B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. VanToai, D.G. Lohnes, J. Chung, and J.E. Specht. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. 39:1464.

Daniel, T.C., A.N. Sharpley, D.R. Edwards, R. Wedepohl, and J.L. Lemunyon. 1994. Minimizing surface water eutrophication from agriculture by phosphorus management. J. Soil Water Conserv. 49:S30–S39.

Erdman, J.W., Jr. 1979. Oilseed phytates: Nutritional implications. J. Am. Oil Chem. Soc. 56:736–741.

Ertl, D.S., K.A. Young, and V. Raboy. 1998. Plant genetic approaches to phosphorus management in agricultural production. J. Environ. Qual. 27:299–304.

Fasoula, V.A., D.K. Harris, and H.R. Boerma. 2004. Validation and designation of quantitative trait loci for seed protein, seed oil, and seed weight from two soybean populations. Crop Sci. 44:1218–1226.

Fehr, W.R. 1991. Principles of cultivar development. Macmillan, New York.

Hulke, B.S., W.R. Fehr, and G.A. Welke. 2004. Agronomic and seed characteristics of soybean with reduced phytate and palmitate. Crop Sci. 44:2027–2032.

Israel, D.W., P. Kwanyuen, and J.W. Burton. 2006. Genetic variability for phytic acid phosphorous and inorganic phosphorous in seeds of soybeans in maturity groups V, VI, and VII. Crop Sci. 46:67–71.

Lei, X.G., and J.M. Porres. 2003. Phytase enzymology, applications, and biotechnology. Biotechnol. Lett. 25:1787–1794.

Li, Z., R.S. Jakkula, R.S. Hussey, J.P. Tamulonis, and H.R. Boerma. 2001. SSR mapping and confirmation of the QTL from PI96354 conditioning soybean resistance to southern root-knot nematode. Theor. Appl. Genet. 103:1167–1173.

Meis, S.J., W.R. Fehr, and S.R. Schnebly. 2003. Seed source effect on field emergence of soybean lines with reduced phytate and raffinose saccharides. Crop Sci. 43:1336–1339.

Nyquist, W.E. 1991. Estimation of heritability and prediction of selection response in plant populations. Crit. Rev. Plant Sci. 10:235–322.

Oltmans, S.E., W.R. Fehr, G.A. Welke, and S.R. Cianzio. 2004. Inheritance of low-phytate phosphorus in soybean. Crop Sci. 44:433–436.

Oltmans, S.E., W.R. Fehr, G.A. Welke, V. Raboy, and K.L. Peterson. 2005. Agronomic and seed traits of soybean lines with low-phytate phosphorus. Crop Sci. 45:593–599.

Pantalone, V.R., F.L. Allen, and D. Landau-Ellis. 2003. Registration of '5601T' soybean. Crop Sci. 43:1123–1124.

Panthee, D., P. Kwanyuen, V.R. Pantalone, C.E. Sams, A.M. Saxton, and D.R. West. 2004. Quantitative trait loci for β-conglycinin (7S) and glycinin (11S) fractions of soybean storage protein. J. Am. Oil Chem. Soc. 81:1005–1012.

Powers, W.J., E.R. Fritz, W.R. Fehr, and R. Angel. 2007. Total and water-soluble phosphorous excretion from swine fed low phytate soybeans. J. Anim. Sci. 84:1907–1915.

Raboy, V. 2002. Progress in breeding low phytate crops. J. Nutr. 32:503S–507S.

Raboy, V., P. Gerbasi, K.A. Young, S. Stoneberg, S.G. Pickett, A.T. Bauman, P.P.N. Murthy, W.F. Sheridan, and D.S. Ertl. 2000. Origin and seed phenotype of maize *low phytic acid 1-1* and *low phytic acid 2-1*. Plant Physiol. 124:355–368.

SAS Institute. 2003. The SAS system for Windows. Release 9.0. SAS Inst., Cary, NC

Spear, J.D., and W.R. Fehr. 2007. Genetic improvement of seedling emergence of soybean lines with low phytate. Crop Sci. 47:1354–1360.

Walker, D.R., A.M. Scaboo, V.R. Pantalone, J.R. Wilcox, and H.R. Boerma. 2006. Genetic mapping of loci associated with seed phytic acid content in CX1834-1-2 soybean. Crop Sci. 46:390–397.

Wilcox, J.R., G.S. Premachandra, K.A. Young, and V. Raboy. 2000. Isolation of high seed inorganic P, low-phytate soybean mutants. Crop Sci. 40:1601–1611.